AFRL-RH-WP-TR-2009-0088

Virtual Coaching Agent for Team Training

Norman I. Badler
Catherine Stocker
Ben Sunshine-Hill

Center for Human Modeling and Simulation
University of Pennsylvania
Philadelphia PA 19104-6389

March 2009

Final Report for December 2007 to March 2009

Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Warfighter Readiness Research Division
Logistics Readiness Branch
Wright-Patterson AFB OH 45433-7604

NOTICE AND SIGNATURE PAGE

AFRL-RH-WP-TR-2009-0088 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNED//

DENNIS J. RIECHMAN
Work Unit Manager
Logistics Readiness Branch

//SIGNED//

DANIEL R. WALKER, Colonel, USAF
Chief, Warfighter Readiness Research Division
Human Effectiveness Directorate
711th Human Performance Wing
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
March 2009	Final	December 2007 – March 2009

4. TITLE AND SUBTITLE
Virtual Coaching Agent for Team Training

5a. CONTRACT NUMBER
FA8650-04-D-6546 DO#15

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62202F

6. AUTHOR(S)
Norman I. Badler, Catherine Stocker, Ben Sunshine-Hill

5d. PROJECT NUMBER
7184

5e. TASK NUMBER
D2

5f. WORK UNIT NUMBER
7184D213

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Center for Human Modeling and Simulation
University of Pennsylvania
Philadelphia PA 19104-6389

8. PERFORMING ORGANIZATION REPORT

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Materiel Command
Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Warfighter Readiness Research Division
Logistics Readiness Branch
Wright-Patterson AFB OH 45433-7604

10. SPONSOR/MONITOR'S ACRONYM(S)
711 HPW/RHAL

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-RH-WP-TR-2009-0088

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
88 ABW/PA cleared on 17Aug2009, 88ABW-09-3693.

14. ABSTRACT
This effort focused on investigating the design and evaluation of an interactive software model to support technical training, as well as for use as an on-the-job task aide. The overarching goals of this research project were to: 1) Utilize motion capture technology to capture the instructional experience of a course instructor for use in the development of a virtual training environment; 2) Provide a mechanism for storing, archiving and efficiently accessing course materials, simulations, Technical Orders (T.O.), multimedia materials, and instructional expertise; 3) Utilize human model avatars as coaches, communicators (instructor surrogates), or simulated maintainers to instruct and illustrate correct and incorrect procedures and practices - especially cautions and warnings; 4) Demonstrate and evaluate a software-based, team training environment for 24/7 instructional access, task preparation, individual practice, and team coordination as an adjunct to hardware trainers and hands-on experience.

15. SUBJECT TERMS Course Resource with Active Materials (CRAM), Interactive Electronic Technical Manuals (IETMs), Avatars, Maintenance Task Training, Virtual Maintainers

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER *(include area code)*
U	U	U	SAR	56	Dennis J. Riechman / NA

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. 239.18

THIS PAGE LEFT INTENTIONALLY BLANK

TABLE OF CONTENTS

ABSTRACT **1**

ACKNOWLEDGEMENTS **2**

1. INTRODUCTION **3**

1.1 Purpose and Scope 3
1.2 Background 4
1.3 CRAM System Overview and Extensions 5
1.3.1 Discussion of CRAM System and Previous Research 5
1.3.2 Addition of Multi-User Capabilities 8

2. METHODS, ASSUMPTIONS, AND PROCEDURES **9**

2.1 General Approach for Evaluating CRAM 9
2.2 Data Collection 9
2.3 Participants 10
2.4 Facilities and Materials 10
2.5 Evaluation Procedures 11

3. RESULTS AND DISCUSSION **13**

3.1 Objective Scores 13
3.2 Subjective Opinions 15
3.3 User Satisfaction 15
3.4 Discussion of Analysis and Evaluation Results 16

4. CONCLUSIONS AND RECOMMENDATIONS **19**

APPENDIX A - CONSENT FORM **21**

APPENDIX B - DEMOGRAPHIC QUESTIONNAIRE **24**

APPENDIX C - CRAM INSTRUCTION SHEETS **27**

APPENDIX D - CRAM OBJECTIVE QUESTIONNAIRE **32**

APPENDIX E - CRAM SUBJECTIVE QUESTIONNAIRES **33**

APPENDIX F - DEBRIEFING STATEMENT **40**

APPENDIX G - SUMMARY OF EVALUATION STATISTICS **42**

APPENDIX H - PARTICIPANT DEMOGRAPHIC DATA **46**

BIBLIOGRAPHY **48**

LIST OF ACRONYOMS **50**

LIST OF FIGURES

Figure 1: CRAM User Interface with Interaction Panels Labeled 6

Figure 2: Assigned Group/Training Method – Mean Objective Scores and 95 percent C.I. 13

Figure 3: “Jacked”/ “Not Jacked” Participants Mean Objective Scores and 95 percent C.I. 14

Figure 4: Participants Feelings Based on Training Method – Mean Scores and 95 percent C.I. 15

Figure 5: Participants Level of Comfort/Agreement with Computers and Technology.. 17

Figure 6: Ranking of Participant's Preferred Training Method...................................... 18

Figure 7: Technology Owned by Participants.. 19

THIS PAGE LEFT INTENTIONALLY BLANK

ABSTRACT

This document reports the results of an evaluation of an interactive, virtual reality system referred to as the Course Resource with Active Materials (CRAM). The evaluation sought to identify the type of familiarization training that would be most useful to students assigned to the Air Force Fighter F-15 Aircraft Maintenance Apprentice Course (J3AQR2A333A025A) at the 82nd Training Wing at Sheppard Air Force Base prior to receiving hands-on training. For purposes of this research, familiarization training focused on allowing students to use the CRAM system to become more aware of the hazards, cautions and warnings involved in accomplishing a maintenance task, when compared to alternative training methods (i.e., reading a Technical Order (T.O.) or watching a video of an instructor performing the task). Aircraft maintenance is a core function performed by numerous personnel in the Air Force. This function encompasses activities such as flightline servicing, equipment repair, as well as the training of new recruits. Various hazards exist within these environments, therefore, performing maintenance tasks correctly is fundamental to trainee safety and equipment integrity. The ability for students to train virtually, prior to hands-on-training, could prevent injuries to personnel as well as excess wear on equipment – all without the oversight of a human maintenance trainer. The results of the evaluation reveal that participants develop an increased awareness of hazards when training with stimulating technology – such as virtual simulations or even videos of an instructor demonstrating a task – versus simply reading from the T.O. The results also indicate a desire by students to train with these technologies over the T.O. Finally, demographic data collected during the evaluation elucidates future directions the Air Force could take to develop a more robust and stimulating virtual training environment to help students become more aware of safety hazards associated with various aircraft maintenance tasks and procedures.

ACKNOWLEDGEMENTS

We would like to express our sincere appreciation to personnel at the 82nd Training Wing (82nd TRW), Sheppard AFB, TX, including Ms. Gina Johnson, as well as instructor and student personnel assigned to the Air Force Fighter F-15 Aircraft Maintenance Apprentice Course (J3AQR2A333A025A), for their outstanding support and participation in this research effort.

1. INTRODUCTION

The "Virtual Coaching Agent for Team Training" research task was sponsored by the Air Force Research Laboratory's Logistics Readiness Branch (711th HPW/RHAL) under the Technology for Agile Combat Support (TACS) contract (FA8650-D-6546, Delivery Order #15). The period of performance for this effort extended from 18 March 2008 to 31 March 2009. Section 1 addresses the scope and purpose of the research effort and provides some background for the context to include summarizing previous research that was accomplished relevant to the field of virtual training. Section 2 explains the methods, assumptions, and procedures associated with the evaluation. Section 3 provides a discussion of the evaluation results, and Section 4 provides conclusions from the research and proposes recommendations for future work. The Appendices contain the complete evaluation plan, along with consent forms and questionnaires.

1.1 Purpose and Scope

This effort represents the last in a sequence of TACS Delivery Orders that focused on investigating the design and evaluation of an interactive software model to support technical training, as well as for use as an on-the-job task aide. The overarching goals of the DO-15 research project were to:

- Utilize motion capture technology to capture the instructional experience of a course instructor for use in the development of a virtual training environment.

- Provide a mechanism for storing, archiving and efficiently accessing course materials, simulations, Technical Orders (T.O.), multimedia materials, and instructional expertise.

- Utilize human model avatars as coaches, communicators (instructor surrogates), or simulated maintainers to instruct and illustrate correct and incorrect procedures and practices - especially cautions and warnings.

- Demonstrate and evaluate a software-based, team training environment for 24/7 instructional access, task preparation, individual practice, and team coordination as an adjunct to hardware trainers and hands-on experience.

Previous research, titled *Extending Interactive Electronic Technical Manuals (IETMs) with Real and Virtual Animated Content for Maintenance Task Training* (AFRL-RH-WP-TR-2009-0027), resulted in the development of a prototype, interactive software system called CRAM (Course Resource with Active Materials). An overview of this system is presented in Section 1.3. The primary purpose of this effort was to demonstrate and evaluate the utility of the CRAM system for supporting aircraft maintenance training in a specific technical training context and to demonstrate the feasibility of simultaneous users interacting to accomplish a multi-person maintenance task.

The CRAM system is intended to support Air Force personnel receiving initial skills training in aircraft maintenance career fields to perform various maintenance procedures. It is intended that any procedure described in an Interactive Electronic Technical Manual (IETM) should be able to be converted into a virtual training procedure. For demonstration and evaluation purposes, the scope of the research effort focused on one particular training event in the F-15 Aircraft Maintenance Apprentice Course conducted at Sheppard AFB, TX, namely, the training of jacking procedures for the F-15 fighter aircraft. The effort included:

a) knowledge acquisition and data collection for a relevant aircraft maintenance training scenario;

b) development of methods and materials to support an evaluation of the CRAM system in the context of chosen training scenario;

c) design and development of CRAM software extensions and enhancements to support multi-user demonstration and evaluation activities; and

d) the analysis and documentation of evaluation results.

1.2 Background

A 2008 Air Education and Training Command (AETC) White Paper *On Learning: The Future of Air Force Education and Training* describes "Millenial" airmen as technologically-savvy, "Digital Natives" who require instantaneous access to knowledge. This paper calls for an educational infrastructure that both recognizes and utilizes the technological aptitude of airmen by modifying current techniques in knowledge management, continuous learning, and precision learning. These modifications include:

- Development of a "dynamic knowledge repository" that is kept up-to-date by "subject matter experts" to guarantee information is quickly disseminated and integrated into training.

- Making knowledge accessible at all times, in all places, in the best media format for a given subject *and student* to both enhance learning and reduce skill decay.

- Ensuring proper collaboration, communication and social networking that can create a supportive practice environment and allow for the sharing of best practices and lessons learned.

One of the primary goals of this research effort was to conduct an evaluation of the utility of an interactive, virtual reality system called CRAM and its potential to augment current aircraft maintenance initial skills training for Air Force personnel. In a 2008 HQ AETC study titled *Airmen and Technology,* it was reported that 97 percent of airmen believed integrating new technologies was important to enhancing training and performance. However, experimental evidence is more ambiguous about the advantages of technology (per se for improving training and performance). Studies show that in certain situations, simulation-based training yields better results than conventional learning methods (Regian,

Shebilske et al., 1992; Washbush and Gosen, 2001). In other situations, simulations must be augmented with real-world instruction to improve their effectiveness (Bell, Kanar et al., 2008). In addition, simulations may be more appropriate for certain populations (Veenman, Prins et al., 2002), and are best used in learning environments that aim to teach *intuitive* understanding rather than rote memorization (Thomas and Hooper, 1991; Swaak and de Jong, 2001). In fact, even the airmen's feelings towards technology, revealed in other questions, proved more complicated than a first glance would suggest: 54 percent of those familiar with avatars felt their use would enable more effective learning. This means that 46 percent did not, which does not suggest as strong a vote of confidence as literature on the Millenial Generation might suggest.

Hence, it is imperative that simulations are thoroughly tested prior to full-scale integration into an training system. One of the goals of this research effort was to answer two questions. First, will using CRAM cause trainees to become more aware of the hazards involved in a maintenance task, or will simpler methods – such as reading the T.O. description or watching a video of the instructor describe and demonstrate the system – be sufficient? Second, which method would trainees *prefer* to practice with? Students with a predisposition for, or incentives to, learning are more likely to seek out educational opportunities, persist in the face of difficulty, have greater cognitive flexibility and retain material longer (Stipek, 1996; Sheldon and Biddle, 1998; Voke, 2002). It is not enough to demonstrate that a training aid *could* assist students if they have no interest in ever practicing with it.

1.3 CRAM System Overview and Extensions

In response to perceived learning needs for aircraft maintainers, we developed a computer-based, interactive system that could engage learners in a virtual physical task characterized by multiple steps, numerous hazards, and multi-person coordinated activities. The software employs simulation, 3D visualization, multi-user support, video clip displays, a text-based communication feature, and a virtual "coach" agent.

1.3.1 Discussion of CRAM System and Previous Research

Through previous research activities, the CRAM system depicted in Figure 1 was developed and demonstrated. Visible in the figure are a number of user interface elements: the Task List Panel, the Step Detail Panel, the Video Panel, the Chat Panel, and the Interactivity Panel.

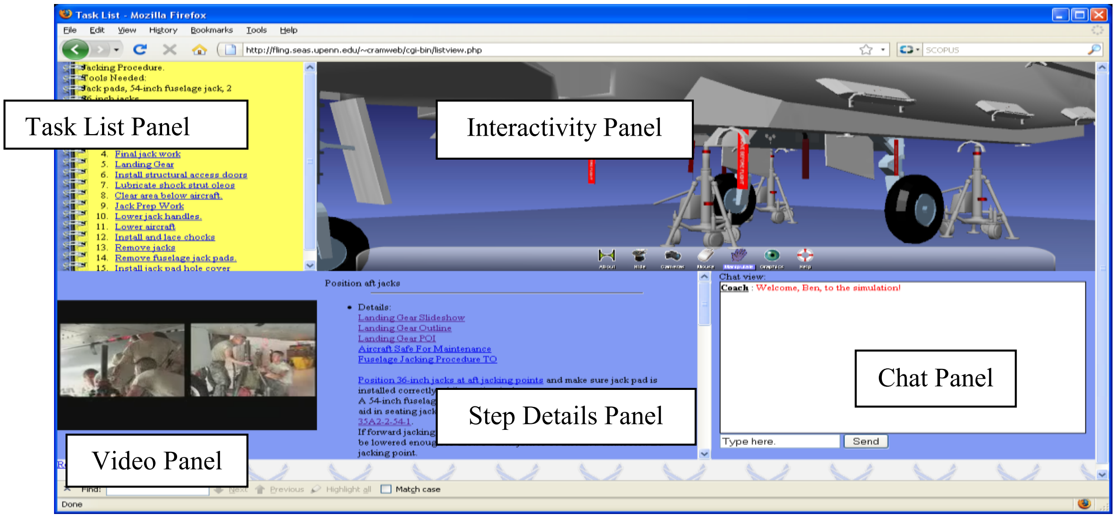


Figure 1: CRAM User Interface with Interaction Panels Labeled

Task List Panel

This panel, located on the upper left side of the interface in Figure 1, displays the top-level sequence of steps in the T.O. The next step which must be completed to advance through the procedure is clearly indicated. Additionally, all steps are clickable; selecting a step causes the details of the step to be displayed in the step detail panel.

Step Details Panel

This panel, located in the bottom center of the interface in Figure 1, displays details of the currently selected step. These details are taken verbatim from the T.O., but may be supplanted with reminders and advice from the instructor, as well as whatever multimedia content is desired. This multimedia, when selected, plays in the video panel. This panel also contains a text box to allow trainees to ask questions and make content suggestions, to be reviewed later by the instructor.

Video Panel

This panel, located to the left of the step detail panel in Figure 1, supports the playback of video and other multimedia content included by the instructor.

Interactivity Panel

This panel, located directly above the step detail panel in Figure 1, displays the interactive 3D world in which the procedure takes place. Trainees can move in the 3D world and manipulate objects using a control scheme familiar to them from 3D video games. During multi-trainee collaboration, other trainees are visible as avatars in this world. Additionally, the Virtual Coach agent appears in the virtual environment when necessary to give advice to the trainee.

Chat Panel

This panel, located in the lower right portion of Figure 1, allows trainees to communicate during multi-trainee collaboration, and its design reflects standard online "instant messaging" behavior. It consists of a text entry field to allow a trainee to type a message, which is then sent to all other trainees participating in the simulation. Past messages sent by all trainees are visible above this field, in chronological order. Additionally, the Virtual Coach uses this space to send his own messages to the trainees when that becomes necessary.

Using CRAM

The design of the CRAM interface is oriented around the concept of providing a trainee various tools to complete a given procedure in the Interactivity Panel. At any time, the user can use the task list panel to monitor his progress through the procedure. The user can click a step to view its' details in the Step Details Panel, but by default that panel displays the details for the next step to be executed. The Step Details Panel is the main conduit for procedure completion information: it contains both the material in the T.O. as well as additional instructional multimedia content. Finally, the Chat Panel is used to communicate with other trainees and the virtual coach.

A trainee may begin the procedure by clicking through the steps in order to see what needs to be done. The trainee then returns to the details of the first step to view any associated video content before beginning. After doing so, the trainee performs the step in the interactive pane. The system, recognizing that the step has been completed properly, advances the simulation and displays the details for the next step. The trainee can proceed to complete that step but has the option of reviewing the details of the first step if desired.

The instructor's interface, when using the CRAM system, is very similar to the trainee's interface with two primary differences. First, questions and comments made by trainees are visible in the instructor's interface either be responded to on an individual basis, or by adding content to the training unit. Second, the instructor has the ability to add content to the details for each task step, which will appear when trainees are viewing or performing the step. This content may consist of text, images, or links to video content or other external multimedia content.

The training content in CRAM can be kept current through instructor's contributions to the wiki website. It is designed to be put online - in a virtual world such as MyBase - making training content instantly and easily available to any airman with internet access. CRAM's framework contains a variety of media (text, video, still images, wiki content, chat and simulation), and has no set time limit for completing training. This allows students to proceed at their own pace, and explore different types of training media suitable to their learning style or preference.

1.3.2 Addition of Multi-User Capabilities

Although using a computer is an inherently single-person activity due to the nature of the input devices, maintaining and repairing an aircraft is not. For instance, in observing the F-15 aircraft jacking procedure, up to eight people were simultaneously involved, each with a distinct role to carry out. There is a significant difference in accomplishing a maintenance procedure alone, versus performing the task as part of a team for the following reasons:

• **Separation of Responsibilities**. A team carrying out a maintenance procedure will assign each participant a role, either explicitly at the outset of the task procedure, or through emergent behavior that occurs during the task. For instance, a participant may decide to work entirely with a single aircraft jack. This decision encourages the participant to think of the procedure in terms of everything happening to that specific jack, which may enable insights that would be lost if the participant were constantly moving to other sites. Of course, only ever working on that jack would limit the trainee's overall exposure to the procedure, meaning that trainees should rotate through roles for maximum benefit.

• **Inter-Participant Communication**. Communcication is key in multi-person tasks, particularly if the task has a specific ordering of required steps, delegated to different participants, and it is not immediately obvious to all participants which steps have, and have not been completed. This is no different in a virtual simulation than in real life. A multi-user training simulation forces participants to communicate with each other, vocally or otherwise, regarding which steps have been completed and which participant is expected to act next. Multi-user integration, therefore, gives trainees exposure to a non-physical dimension of the procedure which would be lost without collaboration.

• **Coordinator Role**. Closely related to the preceding two points, just as a participant may not be directly involved in all physical actions taken during the procedure, so some participants may not physically act at all. This is particularly the case in the jacking procedure, wherein one participant is required to act as a "spotter" to ensure that jacks are raised evenly, reducing the risk of the aircraft falling off the jacks. In a single user environment, there is no need for this role: the participant must run back and forth between jacks anyway, so he himself can step back from the aircraft and assess its attitude on a regular basis. The addition of a multi-user component, therefore, gives trainees experience in a unique maintenance role which would otherwise be ignored.

Some maintenance task procedures require multiple trainees to perform the task, such as jacking an aircraft; however, CRAM does not require that such procedures be carried out with several participants. Flexibility is key: If trainees are encouraged to participate in training simulations on their own schedule, an individual trainee may, at times, wish to participate when a full team of participants is not available. If only a few trainees are available, they should have the ability to double up on roles in order to carry out the entire procedure, even to the point of a single participant fulfilling all roles. Additionally, just as separation of responsibilities gives participants an otherwise unavailable experience, completing a multi-participant procedure as a single trainee offers another perspective.

2. METHODS, ASSUMPTIONS, AND PROCEDURES

2.1 General Approach for Evaluating CRAM

There are various ways to test the knowledge acquired in training simulations – each specific to the type of knowledge trying to be instilled. If the goal is to evaluate memorization, evaluators may track performance time and number of errors along repeated trials in the given virtual task (Gerbaud, Mollet et al., 2008). If the newly acquired skill can only be judged visually, the test may involve expert analysis of the learned movements (Babu, Suma et al., 2007). The CRAM system concentrates on enabling learning by facilitating the uptake of system knowledge. That is, the goal is not to have the trainee memorize the steps of a task procedure or specific physical movements, but to understand the reasoning behind the steps as it relates to the functionality of a system, and to gain an innate understanding of the hazards involved in order to prevent damage to aircraft or personnel. For that reason, the knowledge acquisition of our simulation was tested using a method specifically designed for this evaluation that encompassed showing a set of situational video clips to trainees and asking them to indicate any potential hazards they witnessed, with an explanation of why they perceived the situation to be hazardous or dangerous.

There are also numerous ways to test the usability of a system (Chin, Diehl et al., 1988; Nielsen, 1992; Kirakowski and Corbett, 1993; Nielsen and Landauer, 1993), each with their own advantages and disadvantages (The Federal Aviation Administration, 2009). For purposes of this research effort, we chose to administer the System Usability Scale (SUS), a 10 item Likert scale questionnaire intended to provide a quick and easy method to obtain user satisfaction with the CRAM software (Brooke, 1996). The scale is commonly used to measure the usability of virtual education environments (Babu, Suma et al., 2007; Kaufmann and Dunser, 2007) and has been shown to be one of the more reliable usability questionnaires when used with small sample sizes (Tullis and Stetson, 2004). The SUS questions were added to the subjective questionnaire developed for participants assigned to the CRAM user group which will be discussed further in Sections 2.4 and 2.5. The questionnaire was comprised of a combination of Likert scale questions, essay questions, and rank-ordering questions.

2.2 Data Collection

Data to support extensions to the CRAM system and evaluation was collected over the course of two trips to Sheppard AFB. The initial trip was made to view and videotape students performing the F-15 aircraft jacking procedure, as well as to collect student input on what they found conducive to learning during their training. It was during this trip, through discussion with the students, that the main utility of CRAM became clear. Students repeatedly reported that explanations and stories from instructors were the most memorable and meaningful form of hazard training. In one vivid example, a student mentioned that he did not want to take his wedding ring off during maintenance procedures until an instructor warned that if the ring was caught in the wrong position, a finger could be torn off. This type of informative or experiential explanation of hazards is not included in T.O.s, but was deemed important enough (from a learning standpoint) to incorporate in CRAM for the evaluation.

The videotape from the initial trip was useful for developing a meaningful understanding of the task and requirements of CRAM for content material, but contained only student demonstrations of the task. A DVD was provided that captured a complete demonstration of the aircraft jacking task performed by an instructor, along with useful 3D aircraft models and other course related materials to support the research and evalution. The 3D models were used to create content for both the virtual environment of CRAM, as well as animations used for an objective test conducted as part of the evaluation. The video was segmented into clips that were included in the CRAM system as supplemental information for each task step. The video was also shortened from 1 hour to 35 minutes so it could be used by participants that were assigned to the "video" group during the evaluation.

The purpose of the second trip to Sheppard AFB was to collect expert motion capture data (i.e. an instructor performing or illustrating task steps) that was used to animate the virtual coach in CRAM, and to create the animations portrayed by an avatar used in particular video segments (or scenes) as part of the evaluation. These animated segments demonstrated an avatar correctly (or incorrectly if the motions were purposely altered for the evalaution) performing a particular step in the aircraft jacking task. For the CRAM system, this means that a virtual coach can "step in" and show a user how to complete a step, or warn users of hazardous situations. This is especially useful because, unlike a prerecorded video, students can manipulate their view of the action so that they can see the action performed from the visual perspective of the instructor (a "birds-eye"-view), or from another location in virtual space. For the evaluation, the motion capture was slightly manipulated in two of the three scenes used so that the avatar was performing a task hazard that participants had to identify.

2.3 Participants

Air Force technical training students assigned to the 82nd Training Wing (82 TRW) at Sheppard AFB in the F-15 Fighter Aircraft Maintenance Apprentice Course (J3AQR2A333A025A) represented the population of interest for the evaluation. A total of 48 students, representing six different classes, and up to three different course Blocks (3, 4, and 5), participated in the evaluation. The average age of the participants was 19.9 (range 18-25), and all of the participants were male. Participation did not entail any compensation, and students were made aware that participation was voluntary. The data from one participant was exluded from the analysis after it was discovered that this particular student experienced a problem with the CRAM software that prevented him from completing the evaluation.

2.4 Facilities and Materials

Facilities utilized for the evaluation included training facilities (classrooms and break areas) assigned to the F-15 Aircraft Maintenance Apprentice Course. Participants from each of the six classes were split into three groups and equipment use varied by group. Members of the CRAM condition trained on laptops provided by the investigators, running Microsoft Windows® and preconfigured with the CRAM software. Participants assigned to the video condition viewed a video on a large television screen in a separate area. T.O. condition students were assigned to a separate classroom and reviewed a hardcopy version of the

aircraft jacking procedures. All participants viewed a final video as a group on a large television screen and filled out demographic, objective, and subjective questionnaires.

2.5 Evaluation Procedures

Class sizes participating in the evaluation ranged from seven to nine students per class. Each class was randomly split-up into three groups of approximately the same size (i.e. two to three students per group depending on the class size). The three user groups included the CRAM Group, the Video Group, and the T.O. Group. Prior to the start of the evaluation, all students completed a standard consent form (see Appendix A), and a demographic questionnaire (see Appendix B). After completing the demographic questionnaire, all students were given an instruction sheet (based on the particular group they were assigned to) that described their specific task for the evaluation (see Appendix C). Each group was given a total of 35 minutes to complete their respective task. The time alloted for completing their respective task was derived primarily from the results of pilot testing conducted at the University of Pennsylvania, as well as class scheduling and training constraints at Sheppard.

Participants in the CRAM group used the CRAM system to step through their task in a virtual environment. The CRAM instruction sheet provided brief explanations of how to successfully interact with CRAM, and explained the multi-media content students could explore (short video clips describing a task step and still images of a task step being completed) if they desired. Participants in the video group were instructed to watch a video of an actual instructor describing and demonstrating the F-15 aircraft jacking procedure in a hangar to a class. Participants in the T.O. group were asked to read the description of the F-15 jacking procedure documented in the T.O. All participants were asked to complete their task as if they were preparing to go out to the hangar for actual hands-on, aircraft training at the completion of the evaluation.

After learning the task, participants from all three groups were assembled together in the same room and shown a video containing nine video segments of F-15 aircraft jacking procedures that encompassed both hazardous and non-hazardous events or situations. After each segment was shown, students were asked to document in their objective questionnaire (see Appendix D) what, if any, hazard was prevalent. To ensure that no group had an advantage in demonstrating their intuitive knowledge based on the medium that they trained on, the nine video segments were divded into three different media types:

- Text description of a situation such as the following:
  "After seating the nose and aft jacks (assume they were seated correctly), maintenance personnel raise each jack up (assume all jacking precautions are taken), one by one, in the order they were seated." In this case, the participant would note a hazard occurring because all jacks must be raised *evenly* to maintain aircraft level attitude, ensuring the plane does not slip off the jacks.

- Video clips of steps being performed, such as a clip that may have been altered to portray a maintainer performing a step incorrectly and potentially creating a hazardous situation, or left whole, showing the proper completion of a task step without any potential hazard.

- Animations that have an agent performing a step using motions we acquired by motion capturing an expert (a Staff Sergeant at Sheppard AFB who teaches the F-15 jacking procedure) performing the task. These motions have either been left intact to show the agent demonstrating the step correctly and safely, or have been altered to show the agent performing a step dangerously.

In order to confirm the validity of the expected answers, two experts (maintenance instructors) were also asked to complete the objective questionnaire. The expert responses to questions 5, 6 and 7 (one each: animation, text and live action) did not match the expected answers and it was decided that because those questions may have presented ambigious scenarios or situations, they should be excluded from the evaluation. For example, in one video segment where animation was used, the virtual maintainer was shown raising the ram lock 1 ½ inches and then opening a jack release valve to lower the aircraft. The hazard in this animation is attributed to the fact that the maintainer did not store the jack before lowering the aircraft. Although in the pilot study participants were not noticibly confused by this scenario, neither expert in the evaluation recognized that the plane was being lowered. This may be because of the slow nature in which the plane lowers, as well as confusion over what the virtual maintainer was doing when it was opening the jack release valve. The missing context was apparently important in their understanding of the scenario. Because the actions conveyed in this segment and two others was not clear to the experts, it was deemed unfair to expect the participants to correctly identify the scenes and the hazards in them. Therefore, only questions one through four, as well as questions eight and nine, were considered in the grading of the students' scores and the total possible number of correct answers was six.

After finishing the objective test, participants filled out a subjective questionnaire (see Appendix E) tailored specifically for each of their respective groups that contained some qualitative Likert questions such as the following:

[Practicing a task in CRAM] / [Watching a video of an instructor demonstrating a task] / [Reading the T.O. instructions for the task] could help me become more aware of the hazards involved in a maintenance task.

Some essay questions were also developed to supplement our understanding of what a participant experienced, as well as a question asking participants to rank the three methods of training (CRAM, video, T.O.) in the order they would prefer if they needed to improve their proficiency with a particular task. In addition, the CRAM group had 10 extra questions pertaining to the usability of the system, taken from the System Usability Scale questionnaire (Brooke ,1996). After completing the subjective questionnaire for their respective group, participants were debriefed (see Appendix F) and released back to their instructors.

3. RESULTS AND DISCUSSION

3.1 Objective Scores

A one-way analysis of variance (ANOVA) was calculated on the responses recorded by participants on the objective test, with the independent variable identified as the "assigned training method" used in the evaluation (i.e. CRAM, video, or T.O.). The analysis (see Figure 2) showed that the effect of training method was significant, $F(2,44)=3.50, p< 0.05$. Participants who trained on either CRAM or the video ($M=4.38, SD=1.15$ for both groups) scored higher on the objective test than participants who trained by reading the T.O. ($M=3.40, SD=1.24$).

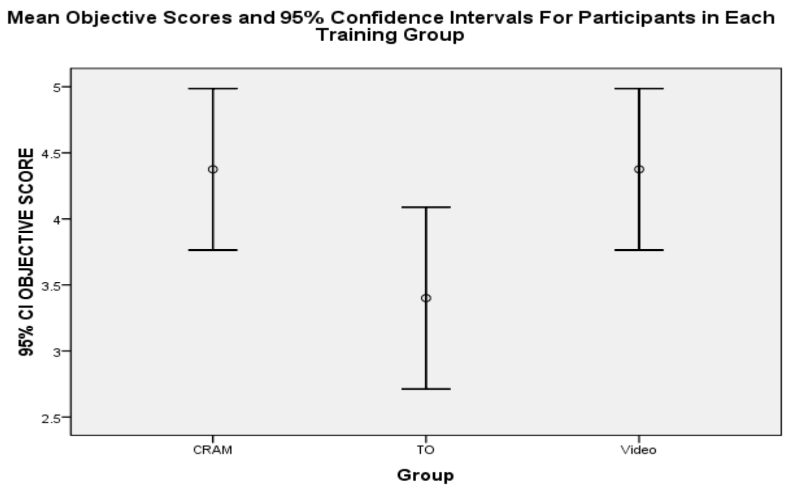


Figure 2: Assigned Group/Training Method - Mean Objective Scores and 95 percent C.I.

A number of additional factors were examined to determine their influence on participants' objective test score. Two of those variables proved significant. First, a one-way ANOVA showed that the effect of previous jacking experience on a participant's objective test score was significant ($F(1,47)=8.566$, $p<0.01$=). Participants who had previous jacking experience ($M=4.57$, $SD=0.992$) scored higher on the objective test than participants who had no previous jacking experience ($M=3.58$, $SD=1.283$). The second variable, the current block of study, produced a significant influence on participants' objective test score ($F(2,44)=6.118$, $p<0.01$). It should be noted that students in Block 5 of the course, had already completed training for the aircraft jacking task, since this training is included in Block 4 of the course Plan of Instruction. Post hoc analysis was done using Tukey's honestly significant difference (HSD) test. Tukey's HSD is a common test used to determine which, among a set of means, are significantly different from the rest. The test indicated that participants in Block 5 ($M=4.57$, $SD=0.992$) performed significantly better on the objective test than participants in Block 3 ($M=3.00$, $SD=1.069$). As depicted in Figure 3, there was no significant effect between either Block 3 or 5 and Block 4 ($M=3.87$, $SD=1.310$).

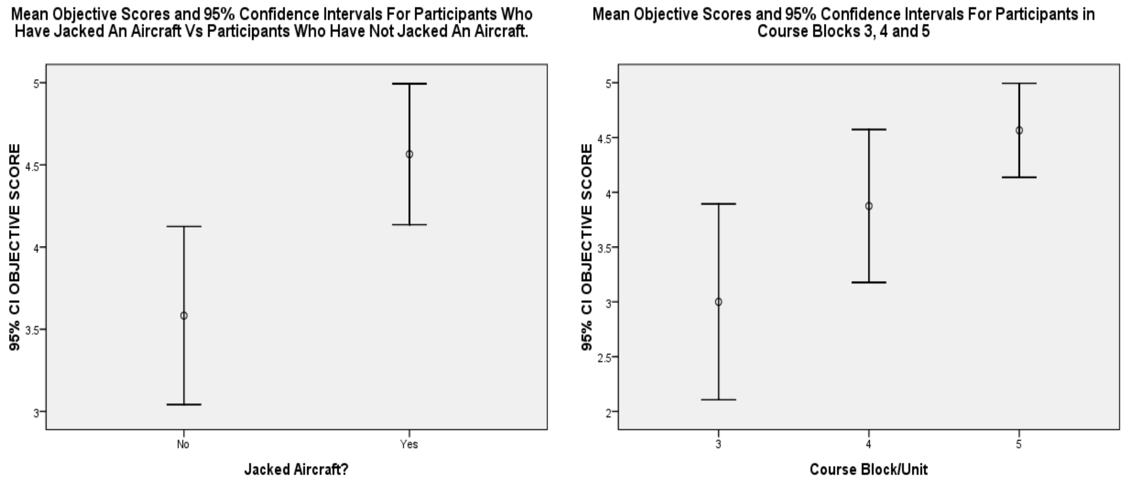


Figure 3: "Jacked"/"Not Jacked" Participants Mean Objective Scores and 95 percent C.I.

Recognizing the influence of jacking experience and training group assignment, a two-way ANOVA was calculated on the objective scores of the participants to test for the interaction of jacking experience and training group assignment. The analysis was significant ($F(2,41)=3.413, p<0.05$) and Post hoc analysis using Tukey's HSD indicated the training group effect on the objective score is greater in the have-not jacked condition than the jacked condition. This means that in the have-not jacked condition, the video and CRAM group did even better against the T.O. than in the have jacked condition.

Other variables from the demographic questionnaire were tested to determine if they had significant effects on the objective score. No effect on objective score was found based on a participant's age ($F(7,39)=0.489, p>0.05$), highest level of education ($F(4,42)=0.371, p>0.05$), rank ($F(3,43)=1.325, p> 0.05$), choice in field assignment ($F(1,45)=0.745, p> 0.05$), or the pairing of a participant with their preferred training method ($F(1,45)=0.503, p>0.05$). In addition, the interaction between technological "inclinations" and training group assignments were tested to check for advantages or disadvantages that the technologically savvy (or inept) participants may have experienced while using the CRAM system. There was no significant effect found on the objective score based on the interactions of training group and a participant's previous experience using interactive computer based training (CBT), ($F(2,41)=0.344, p>0.05$), comfort level with computers ($F(2,38)=0.324, p>0.05$), comfort level with technology ($F(6,35)=2.031, p>0.05$), affinity for learning new technologies ($F(3,38)=0.811, p>0.05$), frequency of video game play ($F(7,33)=2.004, p>0.05$) or frequency of computer usage ($F(3,39)=1.817, p>0.05$).

3.2 Subjective Opinions

In addition to testing the objective improvement in hazard awareness for the three different training groups, the post-training opinions of participants were collected. A one-way ANOVA showed the effect of training method used produced significant opinions in two scenarios. First, when asked to quantify their agreement (1-5, 1=strongly disagree, 5=strongly agree) with the statement that their training method (video or T.O.) better prepared them to jack a real aircraft than virtual training, participants in the video training group (*M=4.07, SD=0.799*) agreed significantly more strongly ($F(1,26)=13.078$, $p<0.01$) than those in the T.O. (*M=2.84, SD=0.987*) training group. In addition, when asked to quantify their agreement (1-5, 1=strongly disagree, 5=strongly agree) with the statement that their training method (CRAM, video or T.O.) would be a good supplement to in-class lecture, participants in the video training group (*M=3.73, SD=0.961*) agreed significantly more strongly ($F(2,42)=5.228$, $p<0.05$) than those in the T.O. (*M=2.36, SD=1.216*) training group. No effect was seen for the CRAM group (see Figure 4).

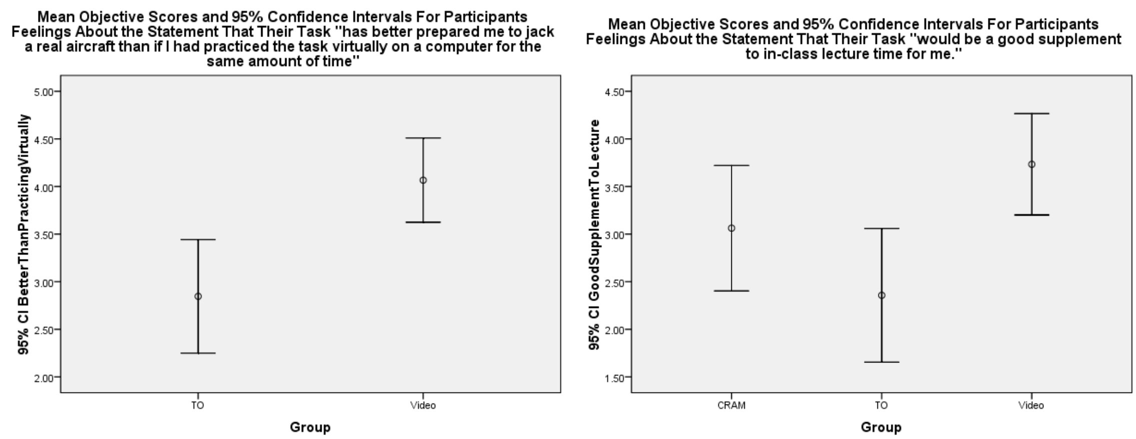


Figure 4: Participants Feelings Based on Training Method - Mean Scores and 95 percent C.I.

Training group assignment did not produce significant results in the subjective situations where participants were asked to: quantify their agreement with the statement that their training method could help them become more aware of hazards ($F(2,42)=0.652$ $p>0.05$), or was better than reading the T.O. ($F(1,29)=1.987$, $p>0.05$) or watching a video ($F(1,27)=0.057$, $p>0.05$).

3.3 User Satisfaction

The user satisfaction with CRAM was measured using the System Usability Scale (SUS). SUS scores range from 0 to 100 and are calculated by subtracting 1 each from the score of odd-numbered questions (1, 3, 5, 7, 9) and subtracting the score of even-numbered questions (2, 4, 6, 8, 10) from 5. Then, the sum of the adjusted scores is multiplied by 2.5 to obtain the final SU value. The average SUS score for CRAM was 60.3. The implications of this score are discussed in the next section.

3.4 Discussion of Analysis and Evaluation Results

The finding that a participant's assigned training group for the evaluation (i.e. CRAM, video, or T.O.) had a significant effect on objective test scores was not unexpected. The initial hypothesis was that training with CRAM could improve hazard awareness over training methods such as reading the T.O. The fact that there is no statistically significant difference in groups that trained on CRAM, or by watching a video, is an interesting finding, but again, not entirely unexpected. Though it is becoming clear that airmen desire training tools that utilize technology, it is not yet known what type of technology, in a given training context, is best suited for training airmen.

It is also not surprising that there is a significant difference between the scores of pariticipants in the "have" and "have-not" jacked aircraft groups. Luckily, this distinction was controlled for by ensuring equal numbers of each training group were in each condition. One point to note is that if retention is this good for the jacking procedure, maybe this task is not the one in most need of external practice methods. The simplest explanation for the difference in scores between participants in different course blocks is that all students of Block 5 had completed the jacking task, while students in Blocks 3 and 4 had not. It is not as obvious why Block 5 is significantly greater than Block 3, but not Block 4. Although not showing a statistical difference does not mean that there is *no* difference, it may mean that there was not a large enough sample, or perhaps the Block 4 subjects had more time to learn about hazards in general or may have informally seen the procedure but not participated in it through formal classwork.

The non significant effects on objective test score by age, level of education, rank, choice in field assignment or pairing of a participant with their preferred training method is reassuring because it suggests that some potential confounding variables had no significant effect. Similarly, with the technology questions, demonstrating no effect between technological aptitude and training performance with CRAM or watching a video suggests that these training methods do not require a high level of technological savvy to operate them.

The subjective results are slightly more interesting in that only the video group expresses significantly stronger agreement over the T.O. when asked if their training method would be a good supplement to in-class lecture. One might suspect that the airmen are afraid or intimidated by the new technology, but looking into the demographic data collected from the airmen, it is evident that this is not the case. For instance, referring to Figure 5, 70.2 percent of participants *strongly* agree with the statement “I am comfortable using a computer” and 25.5 percent more agree less strongly, but do not disagree with the statement. In total, 95.7 percent agree at least to some extent. Similarly, 46.8 percent *strongly* disagree with the statement “I am uncomfortable using technology” and 27.7 percent more disagree less strongly, but do not agree with the statement. In total, 73.5 percent disagree to some extent. Interestingly, even though only 46.8 percent *strongly* disagree that they’re uncomfortable using technology, 59.6 percent *strongly* agree with the statement “I enjoy learning new technologies” and 27.7 percent agree to some extent, totaling 87.3 percent that enjoy learning new technologies.

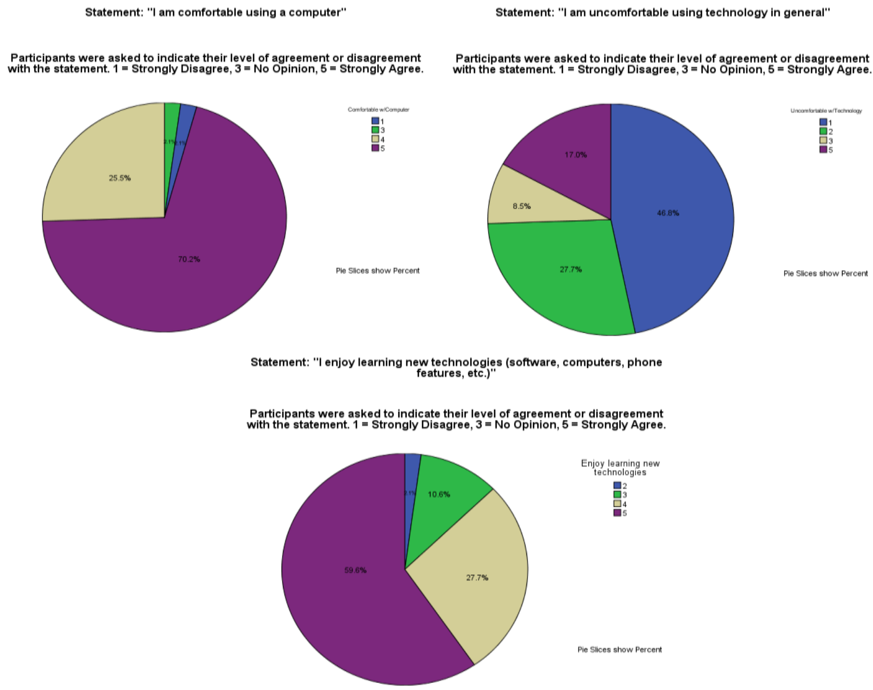


Figure 5: Participant's Level of Comfort/Agreement with Computers and Technology

Not only are the airmen comfortable with technology but based on the results from the demographic questionnaire, they use it to an incredible extent: 91.5 percent of airmen use a computer for social networking, 68.1 percent of them play computer games, 97.9 percent use email and 59.6 percent use some form of instant communication. CRAM not only incorporates all of these features, but teaches hazards as well as a video does. It's possible that the airmen do not realize that something that resembles the video games they play could teach them as well as something drier like a video or reading from a manual.

Although some of the subjective scores point towards participants favoring the video, what should not be discarded is the unfavorable view of studying with the T.O. When asked to put in order their preferred method of practice 91.3 percent chose either virtual training or watching a video, i.e. not the T.O (see Figure 6). Similarly, 76.1 percent chose the other of the two technologies as their second choice. Two-thirds of the participants ranked studying with the T.O. as their last choice. By implementing both training methods (which CRAM does, it has movie clips to go along with virtual steps), 91.3 percent of participants could train on their first choice of training method and would gain a better knowledge of hazards than if they had been asked to study using the T.O.

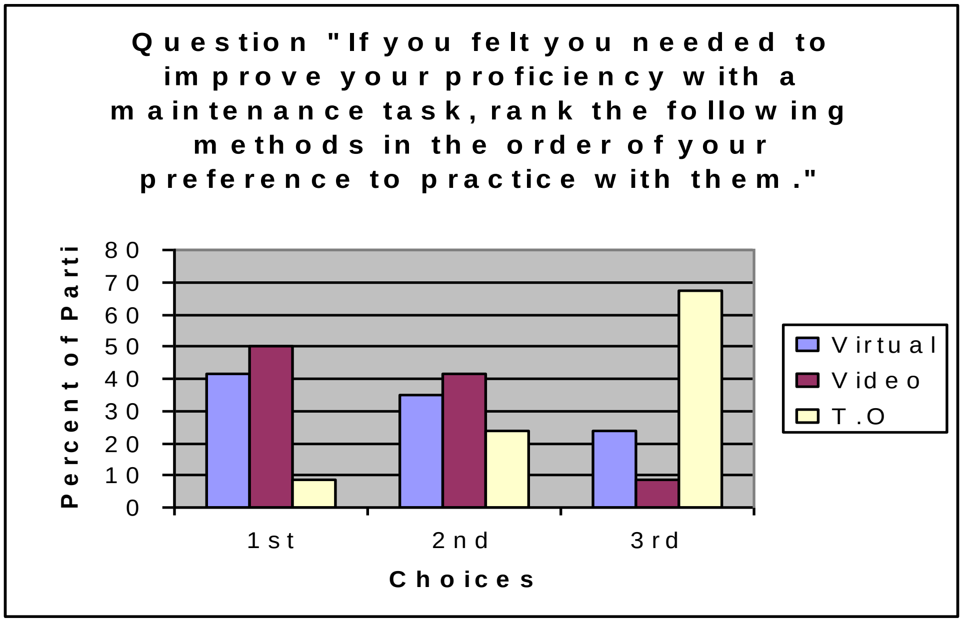


Figure 6: Ranking of Participant's Preferred Training Method

Finally, the average SUS score submitted by CRAM of 60.3 is promising. Scores of 70 and over are generally considered to be the passing rate of usability for a system (Bangor, Kortum et al., 2008), and this score indicates that in only its first iteration as a research prototype, CRAM is almost there. With a few simple improvements, such as those documented by CRAM users in the subjective questionnaires, CRAM could be a very usable system. The improvements are generally simple fixes such as the need for strafe buttons (common in computer games) for better navigation, more guided explanation about what is needed to complete each step – including more demonstration from the virtual coach, and better advertising of all of the components available (video, still images, etc) for each step. Once these improvements are implemented, future versions of CRAM can use the SUS scores collected here as a baseline to confirm or disprove the hypothesis of improvement.

Additionally, though many of the participants recognized in their subjective answers the utility of the CRAM system to allow them to practice safely, the most common complaint from those who did not value the CRAM system as a training tool was that they preferred to practice in hands-on situations. This is to be expected from maintenance trainees, but it may suggest that before having them fill out the SUS questionnaire it should have been more clearly stated that they should answer the questions with the understanding that when they respond to a statement such as “I think that I would like to use this system frequently” (question 1 on SUS) the choice they are making is between this system and another form of training that is not hands-on such as reading the T.O. or watching a video. Based on this choice, as is discussed above, 41.3 percent of participants would prefer CRAM. A summary of all of the analyses can be found in Appendix G, and a summary of the demographic data can be found in Appendix H.

4. CONCLUSIONS AND RECOMMENDATIONS

The purpose of this effort was to demonstrate and evaluate the utility of the CRAM system for supporting aircraft maintenance training in a specific technical training context and to demonstrate the feasibility of simultaneous users interacting to accomplish a multi-person maintenance task. It has been demonstrated that both practicing virtually using CRAM and watching an instructor demonstrate a task in a video can improve an airman's knowledge of hazards significantly over reading about them in the T.O. In addition, when given the choice of what they would like to practice on, participants overwhelmingly chose virtual practice or watching a video over reading the T.O.

This evaluation captured more than just the objective and subjective utility of CRAM; it collected demographic data that gives insight into the future of training. Given the statistics presented in Figure 7, it is evident that rather than presenting CRAM as another CBT system, it should be understood and portrayed as an airman social networking site where trainees can play game-like simulations in which they must avoid hazards (or cause them, in a slightly modified version), can email instructors (through the wiki), and can have instant communication with other airmen during practice simulations, even for multi-person coordinated tasks. Since 74.4 percent of participants are playing video games 1-2 times per week (with 25.5 percent playing daily), that time could be at least partially allocated to playing the CRAM "game." Since 89.4 percent of participants are on the computer 1-2 times per week (with 68.1 percent on daily), if they can get to a computer twice a week, they could practice a task. A very high 93.6 percent of participants have internet access at home, and therefore, any web-based application is within their reach.

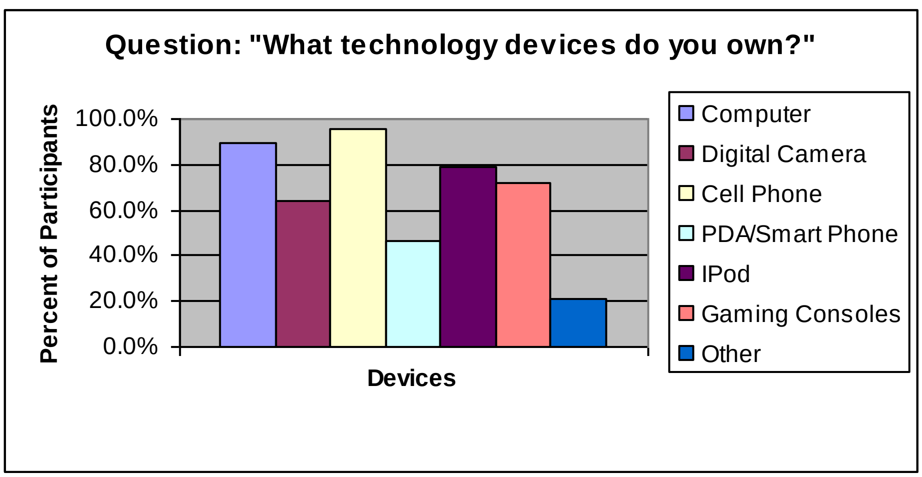


Figure 7. Technology Owned by Participants

It is apparent that any type of technology – even the simplest, such as a video – can be useful for and will be appreciated by maintenance trainees. If it is decided to move forward with videos of instructors demonstrating the task, the next step would be to make those videos even easier to access by making them available on mobile devices. The video capabilities on many cell phones could be utilized (95.7 percent of the participants have cell phones) or they could be made available as podcasts to be viewed on iPods (78.7 percent of participants

have them). If CRAM is to be continued, it could also be made more accessible by being put on a platform that can be accessed by a PDA/Smart Phone (46.8 percent currently) and could be extended to a more game-like environment.

This evaluation was based on the skilled demonstrations of the instructors at Sheppard AFB. Any future expansion of either CRAM or video instructional materials will greatly benefit from careful selection of the best instructional materials and skillful recording and editing of video content. Courses with higher than average washback rates for certain course blocks or training events, and where the practice of procedural steps or tasks in a computer-based system like CRAM might improve training performance, should be investigated as candidates for future technical training research. Among the most promising training events recommended for a follow-on pilot study in the aircraft maintenance technical training arena and identified by instructors during this research effort as a high washback training event in the F-15 course is the aircraft safe for maintenance task. It is a multi-step procedure where the completion of all task steps in the correct sequence is crucial to the successful and safe accomplishment of the the task. Allowing students the opportunity to practice such a task in CRAM (i.e., working outside the classroom/hangar at their own pace) prior to an actual "hands-on" performance evaluation might help reduce the washback rate for this training event.

APPENDIX A - CONSENT FORM

University of Pennsylvania

Center for Human Modeling and Simulation
Norman I. Badler, Professor, Computer and Information Science, 215-898-5862
3330 Walnut St., Philadelphia, PA 19104-6389

Title: Evaluation of the Course Resource with Active Materials (CRAM) System

PURPOSE

The purpose of this research is to assess the effectiveness of different training tools for Air Force maintenance personnel. Should you choose to participate, you will be asked to use one of these methods to learn a task and fill out several questionnaires based on your experience. Your total time is not expected to exceed 90 minutes.

RISKS AND BENEFITS

The potential risks in this project are minimal. There are no direct benefits to you if you choose to participate in this study. However your participation could contribute to a better understanding of Air Force training systems, which could benefit you indirectly and may help future Air Force Trainees.

EXCLUSION CRITERIA

If you are prone to either eyestrain (from use of computer monitors or televisions) or motion sickness, you should not participate in this study.

INJURY/COMPLICATIONS

In the event of any physical injury resulting from the research procedures, medical treatment will be provided without cost to you, but financial compensation is not otherwise available from the University of Pennsylvania or United States Air Force.

COMPENSATION

There is no compensation for being in this study.

CONFIDENTIALITY

Every attempt will be made by the investigators to maintain all information collected in this study strictly confidential, except as may be required by court order or law. Data from the questionnaire answers may be stored on computers but will not be associated with your name. Authorized representatives of the University of Pennsylvania or United States Air Force, including members of the Institutional Review Board (IRB), a committee charged with protecting the rights and welfare of research subjects, may be provided access to research

records that identify you by name. If any publication or presentations results from this research, you will not be identified by name.

WITHDRAWL

Your decision to take part in this study is a voluntary one. You may terminate your participation any time without prejudice to present or future care or services at the University of Pennsylvania or within the United States Air Force.

ALTERNATIVES TO PARTICIPATION

The alternative to being in this study is to not be in this study.

SUBJECT'S RIGHTS

Should you wish further information regarding your rights as a research subject at the University of Pennsylvania, you may contact the Director of Regulatory Affairs, Dr. Sherwin, at 215-898-2614.

By signing below, I assert that:

I have read and understand this consent form.

I am not especially sensitive to either eyestrain (from use of computer monitors or television) or motion sickness. (If so, you should not participate in this study.)

I give permission for the project personnel to use the data collected as a result of this study in any manner they see fit. My anonymity will be maintained unless I give written consent to use my image.

I do not waive any of my legal rights by signing this form.

My signing of this form does not release the investigator, the sponsor, the institution nor its agents from liability for negligence.

______________________________	______________________________
Signature of subject	Signature of person obtaining consent
______________________________	______________________________
Print name of subject	Print name of person obtaining consent

This consent form follows federal regulations. Specifically, Title 45 (Public Welfare), Department Of Health and Human Services, National Institutes Of Health, Office For

Protection From Research Risks, Part 46 (Protection Of Human Subjects). These regulations can be found at http://ohrp.osophs.dhhs.gov/humansubjects/ guidance/45cfr46.htm, specifically sections 46.116 and 46.117.

APPENDIX B - DEMOGRAPHIC QUESTIONNAIRE

Evaluation of the Course Resource with Active Materials (CRAM) System

Participant Number: _____

1. Sex: _____

2. Age: _____

3. Primary Language:________________

4. Please indicate the highest level of education you have completed:
   - ☐ High School
   - ☐ Some College
   - ☐ 2 Year College Degree (Associates)
   - ☐ 4 Year College Degree (BA, BS)
   - ☐ Master's Degree
   - ☐ Other: _________________________

5. What block and unit of instruction are you currently in?

6. Current rank:

7. Have you ever jacked up an aircraft before?

8. Did you choose this career field (upon enlistment or cross training) or were you assigned to this career field in Basic Military Training? Please explain briefly.

9. Have you ever completed an interactive computer-based training (CBT) or soft simulation training course? If so, briefly summarize your experience and the effectiveness of the training.

Please indicate your level of agreement or disagreement with the statements below. 1 = Strongly Disagree, 3 = No Opinion, 5 = Strongly Agree.

	Strongly Disagree			Strongly Agree	
10. I am comfortable using a computer.	①	②	③	④	⑤
11. I am uncomfortable using technology in general.	①	②	③	④	⑤
12. I enjoy learning new technologies (software, computers, phone features, etc.).	①	②	③	④	⑤

13. How frequently do you play video games in an arcade, on a gaming console (Xbox, Nintendo, Play Station, etc) or on a computer, ***that involve virtual environments*** (i.e. computer-simulated worlds that you can move a character through. This excludes games like minesweeper, solitaire and similar puzzle games)?

Every Day	Once or twice a week	Once or twice a month	Rarely	Never

14. How frequently do you use a computer?

Every Day	Once or twice a week	Once or twice a month	Rarely	Never

15. Do you have internet access at home? YesNo

16. If you felt you needed to improve your proficiency with a maintenance task, rank the following methods in the order of your preference to practice with them. 1 = Your first choice, 3 = Your last choice.

Rank	Practice Method
____	Virtual practice of the task on a computer simulated aircraft
____	Watch a video of an instructor describe and demonstrate the task
____	Read the T.O. for the task

17. What do you use the computer for? Please check all that apply.

- ☐ Social Networking (Myspace, Facebook, etc)
- ☐ Computer Games (including games without virtual environments)
- ☐ Work
- ☐ Browsing Websites
- ☐ Email
- ☐ Instant Communication (Skype, instant messaging, etc)
- ☐ Other

18. What technology devices do you own? Please check all that apply.

- ☐ Computer
- ☐ Digital Camera
- ☐ Cell Phone
- ☐ PDA/SmartPhone (iPhone, Blackberry, etc)
- ☐ iPod
- ☐ Gaming Consoles (Xbox, Wii, PS3, etc)
- ☐ Other

APPENDIX C - CRAM INSTRUCTION SHEETS

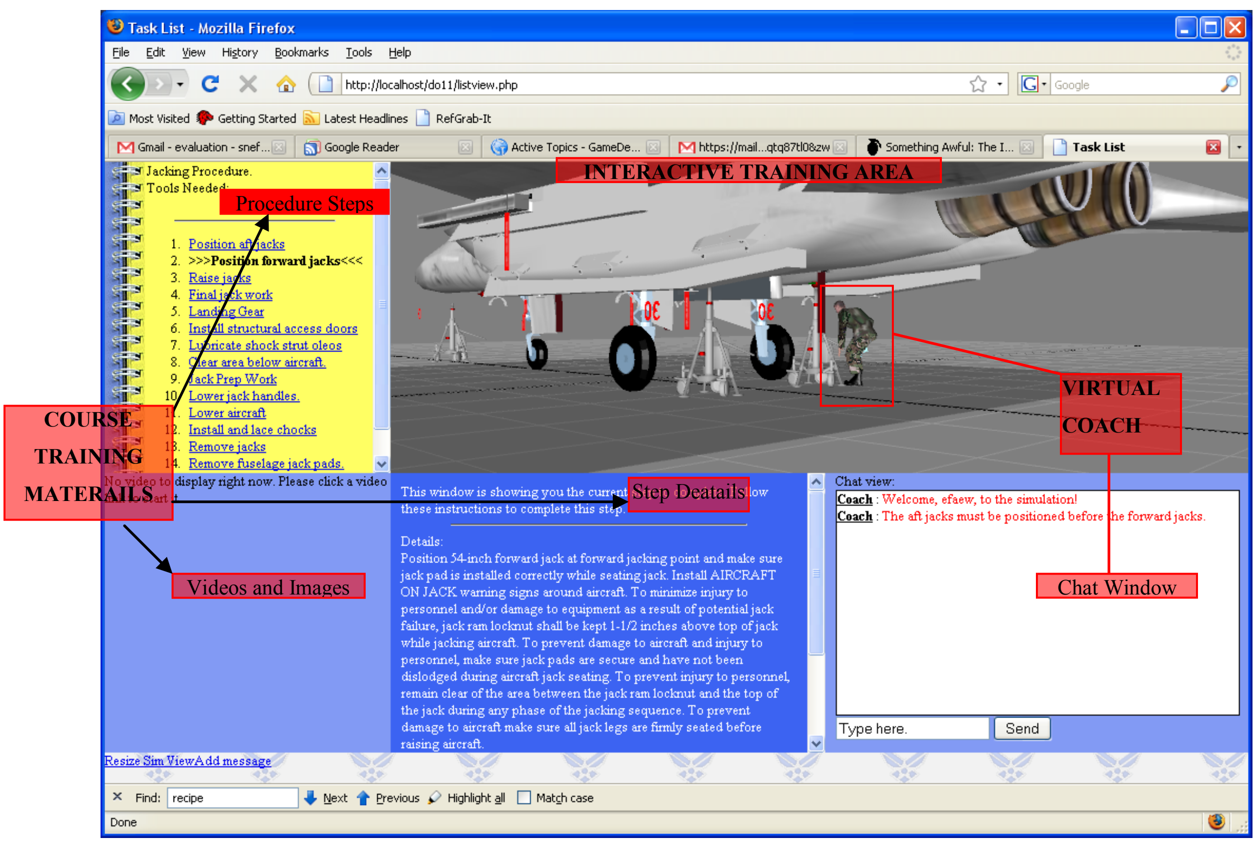
CRAM Instruction Sheet

Thank you for participating in our experiment. Please follow the instructions on this sheet in order; they will help you get oriented in CRAM.

Your task is to use our CRAM system to learn the aircraft jacking task.

Please log-in to the system. To do this, type your participant number into the **username** field and click the “start” button.

Once you are logged in you will see something that looks like the Graphical User Interface (GUI) below. Please note the labels of each area; we will be referring to these sections by their name for the rest of the instruction.



Explanation of each section:

Interactive Training Area: This is the area in the top center of CRAM that contains the virtual representation of the airplane, jacks and coach. You will be asked to move objects in this virtual environment, by using the mouse and dragging the arrows that appear on them.

Virtual Coach: He will pop up in the interactive training area to either inform you when you are making a mistake or warn you of potential dangers. His communication will appear as written text in the chat window at the bottom right of CRAM.

Course Training Materials: The rest of CRAM consists of study materials for your task:

Procedure Steps: Listed in the top left of CRAM, ***you should follow these steps in order, performing each step in the interactive training area*.** The step you have to complete at any time is indicated with arrows. To view the details of any step, click on it, and the details will appear in the step details area.

Videos and Images: Appearing in the bottom left of CRAM for you to review if needed.

Step Details: Further explanation of each step, as described by the T.O., will appear at the bottom center of CRAM. The instructor may additionally include reminders and advice in this area, as well as multimedia content to help you complete the procedure. Multimedia content will appear as underlined links; click the links to view the multimedia content. If you would like to post a question about the task to be answered by an instructor at a later time, you may do so at the bottom of the step details.

Move your mouse over the interactive training area of CRAM. As you pass the mouse pointer over objects which can be moved around, a set of arrows will appear on each side of the object. To move the object, click one of the sets of arrows with your left mouse button and drag it. If no arrows appear, the object cannot be moved at that particular stage in the procedure.

As you pass the mouse pointer over a part which you can interact with at the time, the part will be highlighted. To interact with an object, such as to pump the jack handle, click the left mouse button on the object you wish to interact with. If the part does not highlight when you pass the mouse pointer over it, it cannot be used at that particular stage in the procedure.

To look left and right, first click and release the mouse button in the interactive training window, then use the “A” (look left) and “D” (look right) keys on the keyboard to look around. To move forward and backward, use the “W” (move forward) and “S” (move backward) keys on the keyboard.

You are now ready to begin. You have 35 minutes to complete all of the task steps; please do not speak to any of the other participants during this time. Please study as if you are preparing to go out to the hanger for hands-on training after completion. If you would like, you may follow along in the T.O. as you complete the steps. If you finish within the time provided, feel free to go back over anything you would like to review until the evaluator informs you that it is time to stop.

T.O. Instruction Sheet

Thank you for participating in our experiment. Please follow the instructions on this sheet in order.

Your task is to read the section of the T.O. that describes the aircraft jacking task.

You are now ready to begin. You have 35 minutes to read all of the task steps in the T.O.; please do not speak to any of the other participants during this time. Please study as if you are preparing to go out to the hanger for hands-on training after completion. If you finish within the time provided, feel free to go back over anything you would like to review until the evaluator informs you that it is time to stop.

VIDEO Instruction Sheet

Thank you for participating in our experiment. Please follow the instructions on this sheet in order.

Your task is to use watch a video of an instructor teaching the aircraft jacking task.

You are now ready to begin. You have 35 minutes to watch the movie; please do not speak to any of the other participants during this time. Please study as if you are preparing to go out to the hanger for hands-on training after completion. If you would like, you may follow along in the T.O. as watch.

APPENDIX D - CRAM OBJECTIVE QUESTIONNAIRE

Participant Number:_____

You are now going to watch a short video containing a series of short scenes. Each scene will play once and then you will be asked to answer the question of whether something that happened in the scene would cause a dangerous situation according to what you've learned, or if everything that occurred in the scene was completed safely. Once everyone has completed the question the next scene will begin and you'll be unable to go back to a previous scene. Please be brief and to the point with your answers.

1. Is there a hazard? If so, briefly list all hazards.

2. Is there a hazard? If so, briefly list all hazards.

3. Is there a hazard? If so, briefly list all hazards.

4. Is there a hazard? If so, briefly list all hazards.

5. Is there a hazard? If so, briefly list all hazards.

6. Is there a hazard? If so, briefly list all hazards.

7. Is there a hazard? If so, briefly list all hazards.

8. Is there a hazard? If so, briefly list all hazards.

9. Is there a hazard? If so, briefly list all hazards.

APPENDIX E - CRAM SUBJECTIVE QUESTIONNAIRES

Subjective CRAM Questionnaire

Instructions: Please indicate your level of agreement or disagreement with the statements below. 1 = Strongly Disagree, 3 = No Opinion, 5 = Strongly Agree

Participant Number: _______

System Overview	Strongly Disagree				Strongly Agree
1. Practicing the jacking task in CRAM has better prepared me to jack a real aircraft than if I had **read the T.O. instructions** for the same amount of time.	①	②	③	④	⑤
2. Practicing the jacking task in CRAM has better prepared me to jack a real aircraft than if I had **watched a video of an instructor demonstrating the task** for the same amount of time.	①	②	③	④	⑤
3. Practicing a task in CRAM would be a good supplement to in-class lecture time for me.	①	②	③	④	⑤
4. Practicing a task in CRAM could help me become more aware of the hazards involved in a maintenance task.	①	②	③	④	⑤

5. What did you *like* the most about using the CRAM system?

6. What did you *dislike* the most about using the CRAM system?

7. Do you believe that having access to the CRAM system would be useful during your training? In what way?

8. Was there any particular point(s) during the training session where you felt extremely frustrated – more so than at other points in the session? If so, briefly describe this instance(s) and the reason or cause for frustration.

9. Is there anything specific that you think the CRAM system should do (or not do) to make it more useful?

10. Do you feel more confident after using CRAM to go out to the hangar and perform the aircraft jacking task with your class? If not, what would help you prepare better?

11. Do you feel confident after using CRAM that you know the hazards involved in the task? If not, what would help you prepare better?

System Usability	Strongly Disagree			Strongly Agree	
12. I think that I would like to use this system frequently.	①	②	③	④	⑤
13. I found the system unnecessarily complex.	①	②	③	④	⑤
14. I thought the system was easy to use.	①	②	③	④	⑤
15. I think that I would need the support of a technical person to be able to use this system.	①	②	③	④	⑤
16. I found the various functions in this system were well integrated.	①	②	③	④	⑤
17. I thought there was too much inconsistency in this system.	①	②	③	④	⑤
18. I would imagine that most people would learn to use this system very quickly.	①	②	③	④	⑤

19. I found the system very cumbersome to use. ① ② ③ ④ ⑤

20. I felt very confident using the system. ① ② ③ ④ ⑤

21. I needed to learn a lot of things before I could get going with this system. ① ② ③ ④ ⑤

Subjective Video Questionnaire

Instructions: Please indicate your level of agreement or disagreement with the statements below. 1 = Strongly Disagree, 3 = No Opinion, 5 = Strongly Agree

Participant Number: _______

System Overview	Strongly Disagree				Strongly Agree
1. Watching a video of an instructor demonstrating the jacking task has better prepared me to jack a real aircraft than if I had **read the T.O. instructions** for the same amount of time.	①	②	③	④	⑤
2. Watching a video of an instructor demonstrating the jacking task has better prepared me to jack a real aircraft than if I had **practiced the task virtually on a computer** for the same amount of time.	①	②	③	④	⑤
3. Watching a video of an instructor demonstrating a task would be a good supplement to in-class lecture time for me.	①	②	③	④	⑤
4. Watching a video of an instructor demonstrating a task could help me become more aware of the hazards involved in a maintenance task.	①	②	③	④	⑤

5. What did you *like* the most about watching the video?

6. What did you *dislike* the most about watching the video?

7. Do you believe that having access to videos of an instructor demonstrating tasks would be useful during your training? In what way?

8. Was there any particular point(s) during the video where you felt extremely frustrated – more so than at other points? If so, briefly describe this instance(s) and the reason or cause for frustration.

9. Is there anything specific that you think the video should do (or not do) to make it more useful?

10. Do you feel more confident after watching the video to go out to the hangar and perform the aircraft jacking task with your class? If not, what would help you prepare better?

11. Do you feel confident after watching the video that you know the hazards involved in the task? If not, what would help you prepare better?

Subjective T.O. Questionnaire

Instructions: Please indicate your level of agreement or disagreement with the statements below. 1 = Strongly Disagree, 3 = No Opinion, 5 = Strongly Agree

Participant Number: _______

System Overview	Strongly Disagree				Strongly Agree
1. Reading the T.O. instructions for the task has better prepared me to jack a real aircraft than if I had **watched a video of an instructor demonstrating the jacking task** for the same amount of time.	①	②	③	④	⑤
2. Reading the T.O. instructions for the task has better prepared me to jack a real aircraft than if I had **practiced the task virtually on a computer** for the same amount of time.	①	②	③	④	⑤
3. Reading the T.O. instructions for the task would be a good supplement to in-class lecture time for me.	①	②	③	④	⑤
4. Reading the T.O. instructions for the task could help me become more aware of the hazards involved in a maintenance task.	①	②	③	④	⑤

5. What did you *like* the most about reading the T.O.?

6. What did you *dislike* the most about reading the T.O.?

7. Do you believe that having access to the T.O. outside of class would be useful during your training? In what way?

8. Was there any particular point(s) while reading the T.O. where you felt extremely frustrated – more so than at other points? If so, briefly describe this instance(s) and the reason or cause for frustration.

9. Is there anything specific that you think the T.O. should do (or not do) to make it more useful?

10. Do you feel more confident after reading the T.O. to go out to the hangar and perform the aircraft jacking task with your class? If not, what would help you prepare better?

11. Do you feel confident after reading the T.O. that you know the hazards involved in the task? If not, what would help you prepare better?

APPENDIX F - DEBRIEFING STATEMENT

University of Pennsylvania

Center for Human Modeling and Simulation
Norman I. Badler, Professor, Computer and Information Science, 215-898-7246
200 S. 33rd St., Philadelphia, PA 19104-6389

Evaluation of the Course Resource with Active Materials (CRAM) System

Thank you for your participation in this research on the CRAM System.

Purpose and Hypotheses

The purpose of this research is to evaluate the effectiveness of three different training tools in teaching hazard awareness to Air Force maintenance personnel. The results from this will help us assess the usefulness of the software training tool the University of Pennsylvania created, CRAM. It is expected that students who perform the virtual training exercise within the CRAM software will be better trained on hazard safety, and feel better prepared to perform the aircraft jacking task than students who simply review and familiarize themselves on the procedures, warnings, cautions and hazards by 1) reading a T.O. or 2) watching a video of an instructor performing the procedure. We also anticipate students will *prefer* using CRAM over the other methods to review the material outside of class.

Additional Information

If you would like to know more about training maintenance personnel using virtual worlds, you may be interested in the following articles:

B.S. Bell and A.M. Kanar and S.W.J. Kozlowski. Current issues and future directions in simulation-based training in North America. *In International Journal of Human Resource Management*,19(8), 1416-1434. 2008.
W.L. Johnson and J. Rickel and R. Stiles and A. Munro. Integrating pedagogical agents into virtual environments. *In Presence: Teleoperators and Virtual Environments*, 7(6), 523-546. 1998.

Contact Information

If you have additional questions regarding the research, you may contact Catherine Stocker at cstocker@seas.upenn.edu or Professor Norman I. Badler at badler@seas.upenn.edu. You may keep this document for your records.

Right to Withdraw Data

Your decision to take part in this study is a voluntary one. Your decision whether or not to withdraw your data will not affect present or future care or services at the University of Pennsylvania or the United States Air Force.

By signing below, I assert that:

I have read and understand this debriefing form.

I will not discuss the experiment procedures with any potential participants during the next week.

I give permission for the project personnel to use the data collected as a result of this study in any manner they see fit. My anonymity will be maintained.

I do not waive any of my legal rights by signing this form.

My signing of this form does not release the investigator, the sponsor, the institution nor its agents from liability for negligence.

______________________________	______________________________
Signature of subject	Signature of person obtaining consent
______________________________	______________________________
Print name of subject	Print name of person obtaining consent

APPENDIX G - SUMMARY OF EVALUATION STATISTICS

Independent Demographic Variable *(with dependent variable, objective test score)*	**Significance?** *(*=Approaching, **=Significant)*	**Descriptive Statistics of Objective Scores**	**Explanation**
Training Group (CRAM/Video/T.O.)	**	$F(2,44)=3.496$, $p<0.05$ $(p=0.039)$	Participants in the CRAM and Video groups performed better on the objective test than participants who were in the T.O. group. *CRAM/Video= M=4.38, SD=1.147* *T.O. = M=3.40, SD=1.242*
Jacked Aircraft (Y/N)	**	$F(1,47)=8.566$, $p<0.01$ $(p=0.005)$	Participants that had previously jacked up an aircraft performed better on the objective test than participants who had not previously jacked up an aircraft. *No = M=3.58, SD=1.283* *Yes = M=4.57, SD=0.992*
Course Block Unit (3/4/5)	**	$F(2,44)=6.118$, $p<0.01$ $(p=0.005)$	Participants in Block 5 performed better on the objective test than participants in Block 3. *3 = M=3.00, SD=1.069* *4 = M=3.87, SD=1.310* *5 = M=4.57, SD=0.992* *(5 > 3 by 1.57 p<0.01 (0.004))*
Age (18-25)		$F(7,39)=0.489$, $p>0.05$ $(p= 0.837)$	It cannot be concluded that participant's age affected performance on the objective test.
Highest Level of Education (†)		$F(4,42)=0.371$, $p>0.05$ $(p= 0.828)$	It cannot be concluded that a participant's level of education affected performance on the objective test.
Rank (A1C, AB, Amn,		$F(3,43)=1.325$,	It cannot be concluded that a

Independent Demographic Variable *(with dependent variable, objective test score)*	**Significance?** *(*=Approaching, **=Significant)*	**Descriptive Statistics of Objective Scores**	**Explanation**
Other)		*p>0.05 (p=0.279)*	participant's rank affected performance on the objective test.
Career Field (Chosen/Assigned)		*F(1,45)=0.745, p>0.05 (p=0.393)*	It cannot be concluded that a participant's choice in their field affected performance on the objective test.
Matched First Choice (Y/N)		*F(1,45)=0.503, p>0.05 (p=0.482)*	It cannot be concluded that participant's paired with their first choice affected performance on the objective test.
Jacked Aircraft (Y/N) *x Group* (CRAM/Video/T.O)	**	*F(2,41)=3.413, p<0.05 (p=0.043)*	The group effect on the objective score is greater in the have-not jacked condition than the jacked condition.
Experience Using Interactive CBT (Y/N) *x Group* (CRAM/Video/T.O)		*F(2,41)=0.344, p>0.05 (p=0.711)*	It cannot be concluded that the interaction of a participant's previous experience using interactive CBT's with the type of training they received affected performance on the objective test
Comfortable w/Computer (1-5) *x Group* (CRAM/Video/T.O)		*F(2,38)=0.324, p>0.05 (p=0.725)*	It cannot be concluded that the interaction of a participant's comfort level with computers with the type of training they received affected performance on the objective test
Uncomfortable w/Technology (1-5) *x Group* (CRAM/Video/T.O)	*	*F(6,35)=2.031, p>0.05 (p=0.088)*	The interaction between a participant's comfort level with technology and the type of training they received and its effect on their performance on the objective test approaches significance, but without

Independent Demographic Variable *(with dependent variable, objective test score)*	**Significance?** *(*=Approaching, **=Significant)*	**Descriptive Statistics of Objective Scores**	**Explanation**
			a p-value < 0.05, it can not be concluded to be significant.
Enjoy Learning New Technologies (1-5) *x Group* (CRAM/Video/T.O)		*F(3,38)=0.811, p>0.05 (p=0.496)*	It cannot be concluded that the interaction of a participant's enjoyment in learning new technologies with the type of training they received affected performance on the objective test.
Frequency of Video Game Play (††) *x Group* (CRAM/Video/T.O)	*	*F(7,33)=2.004, p>0.05 (p=0.084)*	The interaction between a participant's frequency of video game play and the type of training they received and its effect on their performance on the objective test approaches significance, but without a p-value < 0.05, it can not be concluded to be significant.
Frequency of Computer Usage (††) *x Group* (CRAM/Video/T.O)		*F(3,39)=1.817, p>0.05 (p=1.435)*	It cannot be concluded that the interaction of a participant's frequency of computer usage with the type of training they received affected performance on the objective test
Better Than Practicing Virtually (1-5)	**	*F(1,26)=13.078, p<0.01 (p=0.001)*	Participants in the Video group felt more strongly than participants in the T.O. group that their method of learning the task had better prepared them to jack a real aircraft than practicing virtually would. *Video= M=4.07, SD=0.799* *T.O. = M=2.84, SD=0.987*

Independent Demographic Variable *(with dependent variable, objective test score)*	**Significance?** *(*=Approaching, **=Significant)*	**Descriptive Statistics of Objective Scores**	**Explanation**
Good Supplement To Lecture	**	$F(2,42)=5.228$, $p<0.01$ ($p=0.009$)	Participants in the Video group felt more strongly than participants in the T.O. group that their method of learning the task was a good supplement to lecture. *Video= M=3.73, SD=0.961* *CRAM = M=3.06, SD=1.237* *T.O. = M=2.36, SD=1.216*
Better Than Reading T.O. (1-5)		$F(1,29)=1.987$, $p>0.05$ ($p=0.170$)	It cannot be concluded that either training group cause participants to feel more strongly that their method of learning the task had better prepared them to jack a real aircraft than reading a T.O. would.
Better Than Watching a Video		$F(1,27)=0.057$, $p>0.05$ ($p=0.813$)	It cannot be concluded that either training group cause participants to feel more strongly that their method of learning the task had better prepared them to jack a real aircraft than watching a video would.
Can Help Hazard Awareness		$F(2,42)=0.652$ $p>0.05$ ($p=0.526$)	It cannot be concluded that participants in any training group felt more strongly that their method of learning helped them to be aware of hazards more than other groups.

† HS, Some College, 2 Yr Degree, 4 Yr Degree, Other

†† The choices for these groups included: daily, once or twice a week, once or twice a month, rarely, or never.

APPENDIX H - PARTICIPANT DEMOGRAPHIC DATA

	Responses	Count	Percentage
Sex	Female	0	0%
	Male	48	100%
Age	18-25	(ave) 19.9	(st dev) 1.6
Primary Language English?	Yes	47	97.9%
	No	1	2.1%
Highest Level Education	HS	29	60.4%
	Some College	14	29.2%
	2 Yr Degree	2	4.2%
	4 Yr Degree	2	4.2%
	Masters	0	0%
	Other	1	2.1%
Course Block Unit	3	8	16.7%
	4	16	33.3%
	5	24	50%
Rank	AB	32	66.7%
	A1C	6	12.5%
	Amn	5	10.4%
	Other	5	10.4%
Jacked Aircraft?	Yes	24	50%
	No	24	50%
Career Field	Chosen	22	45.8%
	Assigned	26	54.2%
Experience Using CBT	Yes	7	14.6%
	No	41	85.4%
Comfortable with Computers	1 (Strongly Disagree)	1	2.1%
	2	0	0%
	3	1	2.1%
	4	13	27.1%
	5 (Strongly Agree)	32	66.7%
Uncomfortable with Technology	1 (Strongly Disagree)	22	45.8%
	2	14	29.2%
	3	4	8.3%

	Responses	Count	Percentage
	4	0	0%
	5 (Strongly Agree)	8	16.7%
Enjoy Learning New Technologies	1 (Strongly Disagree)	0	0%
	2	1	2.1%
	3	5	10.4%
	4	14	29.2%
	5 (Strongly Agree)	28	58.3%
Frequency of Video Game Play	Daily	13	27.1%
	1-2 Times per Week	23	47.9%
	1-2 Times per Month	4	8.3%
	Rarely	6	12.5%
	Never	2	4.2%
Frequency of Computer Usage	Daily	33	68.75%
	1-2 Times per Week	10	20.8%
	1-2 Times per Month	5	10.4%
	Rarely	0	0%
	Never	0	0%
Home Internet Access	Yes	45	93.8%
	No	2	4.2%

BIBLIOGRAPHY

1. Babu, S., E. Suma, et al. (2007). Can Immersive Virtual Humans Teach Social Conversational Protocols? IEEE Virtual Reality 2007 (VR 2007).
2. Bangor, A., P. T. Kortum, et al. (2008). "An Empirical Evaluation of the System Usability Scale." International Journal of Human-Computer Interaction 24(6): 574-594.
3. Bell, B. S., A. M. Kanar, et al. (2008). "Current issues and future directions in simulation-based training in North America." International Journal of Human Resource Management 19(8): 1416-1434.
4. Brooke, J. (1996). SUS: A quick and dirty usability scale. Usability Evaluation in Industry. P. Jordan, B. Thomas, B. Weerdmeester and McClelland. London, Taylor and Francis: 189-194.
5. Chin, J. P., V. A. Diehl, et al. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. ACM Computer Human Interaction '88 (CHI '88), Washington, DC.
6. Gerbaud, S., N. Mollet, et al. (2008). GVT: a platform to create virtual environments for procedural training. IEEE Virtual Reality 2008 (VR '08).
7. HQ AETC. (2008). Presentation: "On Learning: The Future of Air Force Education and Training." Retrieved 2/1/2009, fromhttp://www.aetc.af.mil/shared/media/document/AFD-081216-008.pdf.
8. HQ AETC. (2008). "On Learning: The Future of Air Force Education and Training." Retrieved 2/1/2009, from http://www.aetc.af.mil/shared/media/document/AFD-080130-066.pdf
9. Kaufmann, H. and A. Dunser (2007). Summary of Usability Evaluations of an Educational Augmented Reality Application. HCI International Conference (HCII 2007), Beijing, China.
10. Kirakowski, J. and M. Corbett (1993). "SUMI: the Software Usability Measurement Inventory." British Journal of Educational Technology 24(3): 210-212.
11. Nielsen, J. (1992). Finding usability problems through heuristic evaluation. ACM Computer Human Interaction '92 (CHI '92) Monterey, CA.
12. Nielsen, J. and T. K. Landauer (1993). A Mathematical Model of Finding of Usability Problems. INTERCHI '93, Amsterdam, NL
13. Regian, J. W., W. Shebilske, et al. (1992). "A preliminary empirical evaluation of virtual reality as an instructional medium for visual-spatial tasks." Journal of Communication 42(4): 136-149.
14. Sheldon, K. M. and B. J. Biddle (1998). "Standards, accountability, and school reform: Perils and pitfalls." Teachers College Record 100(1): 164-180.
15. Stipek, D. J. (1996). Motivation and instruction. Handbook of educational psychology D. C. Berliner and R. C. Calfee. New York, MacMillan: 85–113.
16. Swaak, J. and T. de Jong (2001). "Discovery simulations and the assessment of intuitive knowledge." Journal of computer assisted learning 17: 284-294.
17. The Federal Aviation Administration. (2009). "FAA Human Factors Workbench." http://www2.hf.faa.gov/workbenchtools/default.aspx?rPage=ToolList&subCatID=13
18. Thomas, R. and E. Hooper (1991). "Simulations: An opportunity we are missing." Journal of Research on Computing in Education 23(4).

19. Tullis, S. T. and J. N. Stetson (2004). A Comparison of Questionnaires for Assessing Website Usability. Usability Professionals Association (UPA) 2004. Minneapolis, MN.
20. Veenman, M. V. J., F. J. Prins, et al. (2002). "Initial inductive learning in a complex computer simulated environment: The role of metacognitive skills and intellectual ability." Computers in human behavior 18(3): 327-341.
21. Voke, H. (2002). "Student engagement: Motivating students to learn." Infobrief ASCD 2(28).
22. Washbush, J. and J. Gosen (2001). "An exploration of game-derived learning in total enterprise simulations." Simulation & Gaming: An Interdisciplinary Journal 32(3): 281-296.

LIST OF ACRONYOMS

AETC	Air Education and Training Command
AFB	Air Force Base
AFRL	Air Force Research Laboratory
CBT	Computer Based Training
CRAM	Course Resourse with Active Material
IETM	Interactive Electronic Technical Manual
TACS	Technology for Agile Combat Support
T.O.	Technical Order